SEMINARIO DI ANALISI MATEMATICA
DIPARTIMENTO DI MATEMATICA DELL'UNIVERSITA' DI BOLOGNA

P. PLAZZI

INVERSIONE NEL SENSO DI DRAZIN E SISTEMI DIFFERENZIALI DEGENERI

BOLOGNA, 6 DICEMBRE 1984

1. INTRODUZIONE

Nel seguito verranno esposti risultati di esistenza ed unicità relativi al sistema di equazioni ordinarie

(1) $\quad A\dot{x} + Bx = f$

ove $A, B \in C^{nxn}$(C^{hxk} indica l'insieme delle matrici su C con h righe e k colonne: è usuale - in gran parte della letteratura sull'argomento - svolgere la teoria di (1) in riferimento alle matrici, sebbene questa sia essenzialmente relativa agli operatori canonicamente identificati ad esse), f (il termine noto) è una funzione $R \to C^n$ e la incognita $t \to x(t)$ viene ricercata in $C^{(1)}(R,C^n)$: è chiaro che (1) è interessante se A *non* è invertibile.

Nonostante sia relativamente elementare, la (1) merita già un certo interesse per vari motivi:

- sebbene la teoria sia stata elaborata in anni recenti (citiamo in particolare S.L. Campbell) essa ha già raggiunto risultati conclusi, che sono estensioni, per molti versi completamente soddisfacenti, dei risultati e delle tecniche classici (caso di A invertibile);
- al tempo stesso, già nello studio di (1) emergono gli strumenti (in particolare, un tipo generalizzato di inversa per A) che sono i meglio idonei a trattare casi più complessi;
- la possibilità di studiare, sulla base di una esauriente conoscenza di (1), casi più complessi di sistemi degeneri, con le usuali tecniche di estensione: ad esempio, il caso di coefficienti dipendenti dal tempo ([3]), il caso della dimensione infinita ([1], [4]) - per cui tuttavia la teoria è lontana dall'essere approdata a risultati conclusi-

vi-,il caso di equazioni semilineari, il caso dei sistemi d'ordine superiore ([3]);

- il numero notevole di applicazioni di (1) (alcune sono riportate nel testo e nella bibliografia di [3]), che in molti casi hanno preceduto la teoria completa, indice questo sicuro di interesse applicativo.

Il testo seguente è diviso in due paragrafi: nel primo vengono rapidamente esposte le principali proprietà dell'inversa di Drazin di una matrice, che, nell'altro paragrafo, è il principale strumento per studiare la (1).

2. L'INVERSIONE SECONDO DRAZIN

2.1. Per motivare l'introduzione di un nuovo tipo di inversa nel trattare (1), ricordiamo che nel caso 'classico' in cui esiste A^{-1} il problema di integrare (1) è essenzialmente algebrico o, più esattamente, di teoria spettrale (in senso lato) delle matrici; si tratta di calcolare $\exp(A^{-1}Bt)$; poi, moltiplicata la (1) per $\exp(A^{-1}Bt)A^{-1}$, ci si riduce a

$$\frac{d}{dt}(\exp(A^{-1}Bt)x) = \exp(A^{-1}Bt)A^{-1}f$$

risolvibile per integrazione:

$$x(t) = \exp(-A^{-1}Bt)C + \int_0^t \exp(-A^{-1}B(t-s))A^{-1}f(s)ds \tag{2}$$

ove $C = x(o) \in C^n$ è arbitraria.

Questa classica formula suggerisce due considerazioni, utili all'estensione di questi risultati:

a) nella (2) compare in maniera essenziale A^{-1}, accanto all'esponenziale, determinabile concretamente con l'aiuto di proprietà spettrali delle matrici considerate: come si vedrà, l'inversa di Drazin, a differenza di altri tipi di inversa generalizzata, mantiene proprietà siffatte il più vicine possibile a quelle dell'inversione ordinaria;
b) la somma, in (2), riflette il ben noto principio di sovrapposizione che vale anche nel caso degenere: anche qui verrà studiato anzitutto il sistema omogeneo associato a (1),

$$A\dot{x} + Bx = 0 \,. \tag{3}$$

2.2. Come già accennato, i vari processi di inversione generalizzata che consentono di trattare il caso di sistemi degeneri di equazioni di 1° grado, non mantengono in genere le proprietà spettrali dell'inversione usuale, e quelle correlate: ad esempio, se A è invertibile, gli spettri $\sigma(A)$ e $\sigma(A^{-1})$ sono legati da

$$(\sigma(A))^{-1} = \sigma(A^{-1})$$

relazione che *non* vale in genere per autovalori $\neq 0$ se ad A^{-1} si sostituisce l'inversa di Moore-Penrose $A^{\dagger}$ di A (se

$A = \begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix}$ è $A^{\dagger} = \begin{bmatrix} 1/2 & 0 \\ -1/2 & 0 \end{bmatrix}$ e si ha subito $\sigma(A) = \{0,1\}$, $\sigma(A^{\dagger}) = \{0,1/2\}$); questo tipo di proprietà si estende meglio al caso singolare con il tipo di inversione proposto, in un ambiente algebrico più astratto (anelli) da Drazin nel 1958 ([6]): per introdurre la definizione è però opportuno premetterle il

Teorema 1. ([2, p. 121]). *Se* $A \in C^{n\times n}$, *vi è un intero* $k \geq 0$

(detto l'indice di A, *ind* (A)) *che è il minimo intero* ≥ 0 *tale che, equivalentemente:*

i) $N(A^k) = N(A^{k+1})$, *oppure*
ii) $R(A^k) = R(A^{k+1})$;
per questo k *vale la decomposizione* $C^n = R(A^k) \oplus N(A^k)$.

($R(A)$, $N(A)$ indicano l'immagine ed il nucleo di A; $\oplus$ indica come sempre somma diretta).
Il numero ind (A) è una misura della non invertibilità di A, perlomeno nel senso che ind(A) = 0 $\Leftrightarrow$ A è invertibile (si conviene che per B matrice $n \times n$, sia $B^0 = I$ = identità): sebbene enunciato per matrici, il teorema concerne evidentemente l'operatore canonicamente associato ad A, cui si riferisce anche la nozione di indice; con questa osservazione è chiaro come la seguente definizione sia anch'essa intrinseca.

Definizione 1. *Siano* $A \in C^{n\times n}$, k = ind(A). *Si dice inversa di Drazin di* A *una matrice* A^D *tale che*

i) $A^D A A^D = A^D$;

ii) $A^D A = AA^D$;

iii) $A^{k+1} A^D = A^k$.

In effetti, se si indebolisce iii in iii' : *esiste un* $p \in N$ *per cui* $A^{p+1} A^D = A^p$, si ottiene una definizione equivalente a questa: anzi, l'originaria definizione di [6] consta proprio di i-ii-iii' .

La correttezza della definizione si basa sul seguente

<u>Teorema 2.</u> *Ogni matrice* $A \in C^{nxn}$ *ammette un'(unica) inversa di Drazin* A^D.

Per la sua natura costruttiva, riportiamo le grandi linee della dimostrazione (cfr. [2]), marcandole con ■.

■ Con riferimento al teorema 1, è facile vedere che $R(A^k)$, $N(A^k)$ sono sottospazi invarianti per A, la quale sul primo è invertibile mentre è nilpotente (d'indice k) sul secondo. Per la definizione 1, questi sottospazi sono invarianti per ogni eventuale matrice A^D del tipo detto: bisogna che A^D, su $R(A^k)$ inverta A e, su $N(A^k)$, che sia $A^D = 0$: questi requisiti individuano univocamente A^D.■

<u>Osservazioni.</u>1. Se $k = 0$, cioè A è invertibile, è $A^D = A^{-1}$; se $k > 0$ la decomposizione in somma diretta del teorema 1 fornisce per A la rappresentazione

$$(4) \qquad A = P \begin{bmatrix} C & 0 \\ & \\ 0 & N \end{bmatrix} P^{-1}$$

con $P \in C^{nxn}$ invertibile, C quadrata invertibile, N quadrata nilpotente d'indice k, sicché in corrispondenza di (4) la costruzione del teorema 2 dà

$$(5) \qquad A^D = P \begin{bmatrix} C^{-1} & 0 \\ & \\ 0 & 0 \end{bmatrix} P^{-1}$$

che fornisce una rappresentazione a blocchi per A^D.

2. Per come è definita, A^D non è adatta alla risoluzione di

$Ax = b$, cioè, in generale, $A^D b$ non è soluzione di questo sistema, anche se è $b \in R(A)$: ciò avverrebbe se fosse $AA^D A = A$ (se $b = Ac$, allora $A(A^D b) = AA^D Ac = Ac = b$), ma si ha subito che $AA^D A = A$ se e solo se $\text{ind } A \leqq 1$.

■ Se vale $AA^D A = A$ ed $\text{ind}(A) \geqq 1$, con riferimento a (4)-(5) si ha

$$A = AA^D A = P \begin{bmatrix} C & 0 \\ 0 & N \end{bmatrix} \begin{bmatrix} C^{-1} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} C & 0 \\ 0 & N \end{bmatrix} P^{-1} = P \begin{bmatrix} C & 0 \\ 0 & N \end{bmatrix} P^{-1}$$

da cui $N = 0$, cioè $\text{ind } A = 1$; viceversa, trascurando il caso banale $\text{ind}(A) = 0$, se $\text{ind}(A) = 1$, l'equazione da provare coincide con la iii della definizione 1.■

Quindi A^D è un'inversa generalizzata algebrica nel senso di Nashed e Votruba ("A Unified Operator Theory of Generalized Inverses", in [8]) se e solo se $\text{ind}(A) \leqq 1$: in tal caso essa viene in genere chiamata *inversa di gruppo* (cfr. [2]).

2.3. Come si è detto, l'inversione secondo Drazin permette di mantenere per A^D diverse proprietà - specie quelle connesse alle spettrali -, tipiche di A^{-1}: nel seguito menzioniamo alcune di esse, rimandando per la dimostrazione di queste e di altre, a [2].

Teorema 3. *Se* $A \in C^{nxn}$ *e si pone, per* $\lambda \in C$

$$\lambda^{\dagger} = \begin{cases} \lambda^{-1} & \textit{se } \lambda \neq 0 \\ 0 & \textit{se } \lambda = 0 \end{cases}$$

risulta $\lambda \in \sigma(A) \Leftrightarrow \lambda^{\dagger} \in \sigma(A^D)$.

Se $\lambda \in \sigma(A)$, $\lambda \neq 0$, x *è un autovettore generalizzato di grado* p *per* A *relativo a* λ *(cioè* $(A-\lambda I)^p x = 0$, $x \neq 0$), *se e solo se* x *è un autovettore generalizzato di grado* p *per* A^D *relativo a* $\lambda^{\dagger} = \lambda^{-1}$.

La dimostrazione è in sostanza una verifica, sulla base di (4)-(5): per un inquadramento più generale di questo risultato, si veda Ben-Israel e Greville, "Some Topics in Generalized Inverses of Matrices" in [8].

Un'altra proprietà di cui A^D, a differenza di altre inverse generalizzate, gode sempre, e che ha rilevanza anche per il calcolo effettivo di A^D, è la seguente:

Teorema 4. *Se* $A \in C^{n\times n}$, *c'è sempre un polinomio* p *a coefficienti complessi tale che* $A^D = p(A)$.

Come applicazione del teorema 4, consideriamo una proprietà dell'inversione che non si conserva in generale per alcun tipo di inversione generalizzata, la regola del prodotto $(AB)^{-1} = B^{-1}A^{-1}$; essa non vale nemmeno per l'inversa secondo Drazin, anche se in questo caso la formula può essere stabilita sotto la sola ipotesi $AB = BA$, a differenza di altri tipi di inversione:

Teorema 5. *Se* $A, B \in C^{n\times n}$ *commutano, si ha*

$$(AB)^D = A^D B^D = B^D A^D.$$

■ Con la rappresentazione del teorema 4, è facile verificare che A, A^D, B, B^D commutano; si può allora vedere facilmente che $A^D B^D = B^D A^D$ verifica i-ii-iii' nella definizione 1 per $(AB)^D$. ■

Menzioniamo infine una formula di approssimazione per A^D.

Teorema 6. *Se* $A \in C^{n \times n}$, *allora se* $\ell \geqq \operatorname{ind}(A)$,

$$A^D = \lim_{z \to 0} (A^{l+1} + zI)^{-1} A^l .$$

Inoltre si può provare che $0 < \operatorname{ind}(A) = k$ è anche l'ordine della singolarità polare di $\lambda \to R(\lambda, A)$ nell'origine, nel qual caso si ha

$$\lim_{z \to 0} z^k R(z,A) = (AA^D - I)A^{k-1}$$

(cfr. anche [7]).

2.4. Sono state proposte diverse estensioni - sinora non completamente soddisfacenti - dell'inversione secondo Drazin in spazi di Banach qualunque. Nel caso in cui T sia un operatore a indice (nel senso che $\alpha(T) = \delta(T) < +\infty$: cfr. [5]) con polo di ordine $k = \operatorname{ind}(T)$ per $z \to R(z,T)$ nell'origine, si può porre

$$T^D = (2\pi i)^{-1} \int_{|\lambda| = \varepsilon} \lambda^{-1} R(\lambda, T) d\lambda$$

e sviluppare una teoria assai simile a quella in dimensione finita: tuttavia questo impone notevoli restrizioni alle applicazioni. Diverse altre varianti sono state proposte (vedi in particolare Campbell, "The Drazin Inverse of an Operator" in [4]).

3. L'EQUAZIONE $A\dot{x} + Bx = f$, A SINGOLARE

3.1. Cominciamo col considerare alcuni semplici esempi di (1).

A) Siano $A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$, $B = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}$, $f = 0$: (1) diviene

$$\begin{cases} \dot{x}_2 + x_3 = 0 \\ x_1 \qquad = 0 ; \end{cases}$$

si hanno soluzioni solo con dato iniziale $x(o) = \begin{bmatrix} 0 \\ \alpha \\ \beta \end{bmatrix}$, ed in tali casi la soluzione non è mai unica.

B) Siano $A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$, $B = I$: allora il sistema (1) è

$$\begin{cases} \dot{x}_1 + x_1 = f_1 \\ \dot{x}_3 + x_2 = f_2 \\ \qquad x_3 = f_3 \end{cases}$$

ciò richiede $f_3 \in C^{(1)}$: se ciò è vero, $x_2 = f_2 - \dot{f}_3$ e $\dot{x}_1 = -x_1 + f_1$ dà soluzione unica se si impone una condizione iniziale del tipo

$$x(0) = \begin{bmatrix} \alpha \\ f_2(0) - \dot{f}_3(0) \\ f_3(0) \end{bmatrix}, \quad \alpha \in C \text{ arbitrario.}$$

Dunque nel caso che A sia singolare:

1) il dato iniziale è soggetto a vincoli;
2) il termine noto f nella non omogenea deve essere, in generale, abbastanza regolare, se si vuole che il problema di Cauchy

$$(6) \qquad \begin{cases} A\dot{x} + Bx = f \\ x(0) = x_0 \in C^n \end{cases}$$

abbia soluzione;

3) anche nelle ipotesi 2 la soluzione può non essere unica.

Nel seguito tratteremo il caso del problema (6) per la (1) considerando solo soluzioni 'classiche' $x \in C^{(1)}(R, C^n)$: del resto in dimensione finita le stesse ipotesi che garantiscono esistenza ed unicità implicano la possibilità di prolungare soluzioni locali (in dimensione infinita le cose stanno ben altrimenti: per tutta la questione cfr. Campbell cit., in [4]). Non stupisce dunque il prevalere dell'aspetto algebrico nella trattazione di (6), come nel caso di A invertibile.

3.2. Introduciamo la seguente terminologia:

Definizione 2. x_0 *si dice un vettore (o condizione) iniziale compatibile con* (1) *se il problema* (6) *ha (almeno una) soluzione; la* (1) *si dice trattabile se* (6) *ha soluzione unica per ogni condizione iniziale compatibile.*

Nel seguito considereremo soprattutto la trattabilità di (1): seguendo un procedimento euristico classico, si può applicare in (6) una trasformazione (formale) di Laplace:

$$L(A\dot{x}) + L(Bx) = L(f)$$

da cui

$$(sA + B)L(x) = L(f) + Ax_0$$

che suggerisce una stretta connessione tra unicità ed esistenza del risolvente generalizzato $s \to (sA + B)^{-1}$; in effetti vale il

Teorema 7. ([3, p. 34]). *Se* $A, B \in C^{n \times n}$ *l'equazione omogenea*

(3) *è trattabile se e solo se:*

(7) $\exists \lambda \in C$ tale che $\lambda A + B$ è invertibile .

Poiché questo teorema è la chiave della trattazione di (1) in [3], diamo le grandi linee della dimostrazione .

■ Valga (7): definiamo le matrici $\hat{A} = (\lambda A + B)^{-1}A, \hat{B}=(\lambda A+B)^{-1}B$; poiché $\lambda\hat{A} + \hat{B} = I$, $\hat{A}$ e $\hat{B}$ commutano e (3) equivale a $\hat{A}\dot{x} + \hat{B}x = 0$; a meno di similitudini si può scrivere

$$\hat{A} = \begin{bmatrix} C & 0 \\ 0 & N \end{bmatrix}$$ con C invertibile, N nilpotente, da cui

$$\hat{B} = \begin{bmatrix} I - \lambda C, & 0 \\ 0 \quad , & I-\lambda N \end{bmatrix} = \begin{bmatrix} \hat{B}_1 & 0 \\ 0 & \hat{B}_2 \end{bmatrix} \quad ; \text{ posto } x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

(3) si spezza nel sistema

$$\begin{cases} C\dot{x}_1 + (I-\lambda C)x_1 = 0 \\ N\dot{x}_2 + (I-\lambda N)x_2 = 0 \ ; \end{cases}$$

poiché la prima equazione è chiaramente trattabile, basterà provare che tale è la seconda; a questo scopo proviamo che ha la sola soluzione $x_2 = 0$. In effetti, se $k = \text{ind}(N)$, applicando N^{k-1} segue ($(I - \lambda N)$ è invertibile) $N^{k-1}x_2 = 0$; quindi $N^{k-1}\dot{x}_2 = 0$ e, applicando N^{k-2}, $N^{k-2}\dot{x}_2 = 0$ e così via.

Viceversa, (3) sia trattabile e, per assurdo,

$$\forall \lambda \in C \ \exists \ v_\lambda \in C^n, \ v_\lambda \neq 0$$

$$\text{tale che } (\lambda A + B)\, v_\lambda = 0 .$$

Si proverà che in tale ipotesi si può costruire una soluzione non nulla del problema (3) con condizione $x(0) = 0$, ciò che contraddice la trattabilità di (3).

Una tale soluzione si costruisce così: se $\lambda_1,\ldots,\lambda_r (Re\lambda_1 < \ldots < Re\lambda_r)$ sono tali che i corrispondenti v_λ sono linearmente dipendenti, si scelgano scalari non tutti nulli $\alpha_1 \ldots \alpha_r$ per cui

$$\sum_{j=1}^{r} \alpha_j v_{\lambda_j} = 0; \text{ allora } z(t) = \sum_{j=1}^{r} \alpha_j \exp(\lambda_j t) v_j \text{ è la}$$

soluzione cercata. ■

Dal teorema 7 segue una caratterizzazione completa della consistenza e della trattabilità per l'omogenea (3).

Teorema 8. ([3]). *Se* (3) *è trattabile, cioè vale* (7), *allora*

i) *un* $x_0 \in C^n$ *è consistente per* (3) *se e solo se*

$$x_0 \in R(\hat{A}^D \hat{A});$$

ii) *l'integrale generale di* (3) *è dato da*

$$x(t) = \exp(-\hat{A}^D \hat{B} t)\, \hat{A}^D \hat{A} q, \quad q \in C^n. \tag{8}$$

Nell'enunciato si sono seguite le notazioni introdotte nella prova del teorema 7: si può peraltro dimostrare che $\hat{A}^D\hat{A}$, $\hat{A}^D\hat{B}$ non dipendono da λ. Si noti ancora l'analogia formale tra (8) e la (2) nel caso omogeneo.

3.3. Per l'equazione non omogenea (1), oltre alla (7), bisogna, come si è visto, fare alcune ipotesi sulla regolarità di f. Si ha

Teorema 9. ([3]). *Si consideri l'equazione* (1), *nell'ipotesi* (7); *di più, posto* $k = \mathrm{ind}(\hat{A})$, *sia* $f \in C^{(k)}(R, C^n)$. *Allora la* (1) *ha sempre soluzione, una essendo data da*

$$x_1(t) = \exp(-\hat{A}^D\hat{B}t)\int_0^t \exp(\hat{A}^D\hat{B}s)\,\hat{A}^D\hat{f}(s)ds +$$

$$+ (I-\hat{A}^D\hat{A})\sum_{i=0}^{k-1}(-1)^i(\hat{A}^D\hat{B})^i\hat{B}^D\hat{f}^{(i)}(t) = x_0(t) + r(t).$$

Posto $\hat{w} = r(0)$, *si ha:*

i) $x_0 \in C^n$ *è consistente per* (1) $\Leftrightarrow$ $x_0 \in \hat{w} + R(\hat{A}^D\hat{A})$;

ii) *la* (1) *è trattabile e l'unica soluzione di* (6) *con condizione iniziale consistente* $x(o) = x_0$ *è*

$$x(t) = \exp(-\hat{A}^D\hat{B}t)\hat{A}^D\hat{A}\,x_0 + x_1(t). \tag{9}$$

Anche in questo enunciato si sono seguite le notazioni precedenti, ponendo di più $\hat{f} = (\lambda A+B)^{-1}f$; ancora, k, $\hat{A}^D\hat{B}$, $\hat{A}^D\hat{f}$, $\hat{A}^D\hat{A}$, $\hat{B}^D\hat{f}$, $\hat{w}$ sono indipendenti da λ.

Nella (9) è evidente l'analogia con la (2): si noti che il termine $r(t)$ che non compare nel caso classico, scompare per A invertibile perché $\hat{A}^D\hat{A} = I$.

Menzioniamo infine a questo punto il caso di sistemi rettangolari ($A, B \in C^{n\times m}$ con $n \neq m$), sovradeterminati o sottodeterminati: essi si possono trattare secondo questa falsariga, se alla condizione (7)

si sostituisce una ipotesi di iniettività (rispettivamente, suriettività) per una matrice $(\lambda A + B)$; rimandiamo a [2] o a [3].

BIBLIOGRAFIA

[1] CAMPBELL, S.L.,"The Drazin Inverse of an Infinite Matrix", SIAM J. Appl. Math. 31(1976), 492-503.

[2] CAMPBELL, S.L. e MEYER, C.D., Jr. , Generalized Inverses of Linear Transformations, "Surveys and Reference Works in Mathematics" 4, Pitman Ed. 1979.

[3] CAMPBELL, S.L.,Singular Systems of Differential Equations ,"Research Notes in Mathematics" 40,Pitman Ed. 1980.

[4] CAMPBELL, S.L. (Ed.), Recent Applications of Generalized Inverses , "Research Notes in Mathematics" 66,Pitman Ed. 1982.

[5] DOWSON, H.R., Spectral Theory of Linear Operators, "London Mathematical Society Monographs" 12, Academic Press 1978.

[6] DRAZIN, M.P., "Pseudoinverses in Associative rings and semigroups", Am. Math. Monthly 65 (1958), 506-514.

[7] KATO, T., Perturbation Theory for Linear Operators ,"Die Grundlehren der mathematischen Wissenschaften in Einzeldarstellungen" 132, Springer Verlag 1976.

[8] NASHED, M.Z. (Ed.), Generalized Inverses and Applications,Academic Press 1976.